1.0

2.8 2.5

3.2 2.2

3.6

4.0 2.0

1.1

1.8

1.25 1.4 1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

A LINEAR THEORY FOR NONCAUSALITY

BY

J. P. FLORENS and M. MOUCHART

TECHNICAL REPORT NO. 51
NOVEMBER 1981

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH
THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

82 06 09 021

A LINEAR THEORY FOR NONCAUSALITY

J. P. Florens
CERAMMISE, University of Aix-Marseille
Visiting Stanford University

and

M. Mouchart
CORE, University of Louvain

TECHNICAL REPORT NO. 51

NOVEMBER 1981

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

A LINEAR THEORY FOR NON-CAUSALITY

J.P. Florens
CERAMMISE, University of Aix-Marseille
Visiting Stanford University

and

M. Mouchart
CORE, University of Louvain

ABSTRACT

Different definitions of non-causality (according to Granger, Sims, Haugh and Pierce,...) are analyzed in terms of orthogonality in the Hilbert space of square integrable variables. Conditions, when necessary, are given for their respective equivalence. Some problems of testability are mentioned. Finally non-causality is also analyzed in terms of "rational expectations", extending previous results of Sims.

A LINEAR THEORY FOR NON-CAUSALITY

J. P. Florens and M. Mouchart

1.1. Introduction.

Following Granger's (1969) and Sims' (1972) papers, the non-causality concept has taken on great importance in econometrics literature. This concept is essentially the same as the concept of transitivity introduced into statistics by Bahadur (1954) and used in sequential analysis (see e.g. Hall, Wijsman and Gosh (1965)). Intuitively, transitivity can be presented in the following way: a sub-process $(z_n)_n$ of a multivariate stochastic process $(x_n)_n$ is transitive if the past and current values of z_n are sufficient to forecast z_{n+1}. Equivalently, if x_n is partitioned into (z_n, y_n), we say that the process generating y_n does not cause the process generating z_n.

A precise statement of this intuitive definition can be made in different ways. In some of our previous work, non-causality is couched in terms of sequences of independence conditions between σ-fields (see Florens, Mouchart, 1980a,b and Florens, Mouchart, Rolin, 1980). In this paper we propose definitions in terms of sequences of orthogonality conditions between linear subspaces of the Hilbert space of random variables. This kind of presentation is implicit in most of econometrics papers and was explicitly used by Hosoya (1977). (In economics or econometrics literature, see also e.g. Gourieroux, Montfort (1980) or Futia (1981) in which the same kind of mathematical tools are used. In time series literature this kind of presentation is very common. See e.g. Anderson (1971)).

1

The main purpose of this paper is to show the equivalence of several definitions of non-causality (Granger (1969), Sims (1972, 1980), Haugh and Pierce (1977)). These authors have often simultaneously given a definition of and a test procedure for non-causality. We essentially analyze here the relations between these definitions. Comparison of the properties of test procedures is clearly another story. For example, we shall never use stationarity assumptions in the definitions of non-causality or in the proofs of their equivalence. However stationarity is a crucial assumption in test procedures.

An important point about non-causality is its relationship with the exogeneity concept used in econometrics literature or, more generally, with the theory of sufficiency and ancillarity in sequential models. This relationship was the main topic of our previous papers (Florens, Mouchart 1980a, Florens, Mouchart, Rolin 1980) in any of which a bibliography can be found. In particular, the relationship with exogeneity is studied in Florens, Mouchart (1980c) and in a paper by Engle, Hendry, Richard (1980). So this point will not be treated here.

This paper is organized in the following way. Notation is presented in the second part of the introduction. Section 2 is devoted to the original definition of Granger and to the main properties of this concept. Sims' first definition, Haugh, Pierce's definition and their respective equivalences to Granger's definition are given in Sections 3 and 4. In Section 5 links between non-causality and rational expectations are pointed out and Sims' second definition is presented and shown to be equivalent to Granger's. Definitions, notation and results on orthogonality in a Hilbert space are recalled in the appendix.

2

1.2. Notations.

Let $(\Omega, \mathcal{O}, P)$ be a probability space and L^2 be the Hilbert space of square integrable random variables (defined P - almost surely). In this paper inner product, orthogonality, projection, completion... are relative to the canonical structure of L^2 (see any book on probability theory e.g. Neveu (1964)). For simplicity we restrict our presentation to a bivariate discrete stochastic process $(x_n)_n = (y_n, z_n)_n$ $n=0,1...$ i.e. to a double sequence of random variables $(y_n)_n$ and $(z_n)_n$. All random variables considered are assumed to be elements of L^2. (Note that random variables are defined only almost surely, so we in fact consider a class of stochastic processes such that each is a modification of others).

It must be pointed out that the time index belongs in $\mathbb{N} = \{0,1,...\}$ and not in $\mathbb{Z} = \{...,-1,0,1,...\}$. (Continuous time is another story!)

This hypothesis does not limit our results and has the advantage of making clear the scale of initial conditions. In fact $\mathbb{N}$ must be completed with a maximum element ∞. If $\mathbb{Z}$ is the time index, it must be completed by a minimum element $-\infty$ and a maximum element ∞. So $-\infty$ if the index time is $\mathbb{Z}$, and 0 if the index time is $\mathbb{N}$, play the same role, (in Florens-Mouchart (1980 b) details can be found about this modification of the time index).

Let $(y_n)_{n=0,1,...}$ be a stochastic process. We denote by y_n^m $(n \le m)$ the linear subspace of L^2 generated by $y_n,...y_m$. (For example, y_n^n is the subspace generated by y_n). If m is finite, such finite-dimensional subspaces are closed. y_n^∞ denotes the closed subspace generated by $y_n, y_{n+1},...$. y_0^n represents the history (in the sense of all linear functions of the past) of y_{n+1} . Similar notation is used for $(z_n)_n$.

2. Granger's Non-Causality

For expository purposes, we first recall Granger's (1969) concept of non-causality along with some of its main properties.

Definition 2.1. y does not linearly cause z iff

$$(2.1) \qquad \forall \, n \geq 0: \, z_{n+1} \perp y_0^n \mid z_0^n + u \qquad \square$$

In this definition, u may represent initial conditions and any other relevant information to be used by means of linear functions. Typically u will include, at least, the constant functions, and also any information available at the start of the process. Information that becomes available later will be introduced in section 5. u is a closed linear subspace of L^2.

For instance, when u contains the constant functions only, u may be dropped in definition 1 if either of the processes y or z have zero-mean. From the definition of conditional orthogonality (see Theorem A.1 in the appendix), condition (2.1) allows several readings. The projection of z_{n+1} onto the linear space $y_0^n + z_0^n + u$ is contained in the linear space $z_0^n + u$. Alternatively, the residual of the projection of z_{n+1} onto the linear space $z_0^n + u$ is orthogonal to the linear space y_0^n.

By theorems A.4 and A.8, (2.1) is equivalent to any one of the following properties:

$$(2.2) \qquad \forall \, n \geq 0: \, z_0^{n+1} \perp y_0^n \mid z_0^n + u$$

4

(2.3) $\quad\quad\quad$ $\forall\,(n,p),\ 0 \le p \le n:\ z_{n+1} \perp y_p \mid z_0^n + u$

(2.4) $\quad\quad\quad$ $\forall\,(n,p),\ 0 \le p \le n:\ z_0^{n+1} \perp y_0^p \mid z_0^n + u$

If (2.1) – or (2.2) or (2.3) or (2.4) – is not satisfied, we shall say that "<u>y linearly causes z</u>". Linear causality and non-causality enjoy several interesting properties.

<u>Property 2.2</u>. For any process z, z does not linearly cause z.

Indeed, by theorem A1,

(2.5) $\quad\quad$ $\forall\,u$ $\quad\quad$ $\forall\,n \ge 0:\ z_{n+1} \perp z_0^n \mid z_0^n + u$.

<u>Property 2.3</u>. Linear causality is not transitive.

In other words, y linearly causes z and z linearly causes w do <u>not</u> together imply y linearly causes w. (Note that "transitivity" is used here in the usual algebraic sense).

We shall call "y and z" the set of processes $(\alpha y_n + \beta z_n)_n$ for any α and β in $\mathbb{R}$. This set may also be viewed as the linear space generated by y_n and z_n. The history of "y and z" up to the instant n is defined as $y_0^n + z_0^n$. In other words, "y and z" represents the set of information obtainable <u>linearly</u> from the observations y_n and z_n.

Property 2.4. y and z does not linearly cause w if and only if neither y nor z linearly causes w.

In other words, and this is a direct implication of theorem A.3, one has

$$(2.6) \qquad \forall\, n \geq 0: \quad w_{n+1} \perp (y_0^n + z_0^n) \mid w_0^n + u$$

if and only if

$$w_{n+1} \perp y_0^n \mid w_0^n + u \qquad \forall\, n \geq 0$$

and

$$w_{n+1} \perp z_0^n \mid w_0^n + u \qquad \forall\, n \geq 0$$

Property 2.5. y does not linearly cause z and w does <u>not</u> imply that y does not linearly cause z (or that y does not linearly cause w).

It is therefore possible that y linearly causes z, y linearly causes w but that y does not linearly cause z and w. In other words:

$$(2.7) \qquad \forall\, n \geq 0 \,, \quad \forall\, (\alpha,\beta) \in R^2 :$$

$$(\alpha z_{n+1} + \beta w_{n+1}) \perp y_0^n \mid z_0^n + w_0^n + u$$

does not imply:

$$z_{n+1} \perp y_0^n \mid z_0^n + u \,.$$

6

Property 2.6. y linearly causes z and w does not imply that
either y linearly causes z or that y linearly causes w.

It is therefore possible that y linearly causes z and w and
that neither y linearly causes z nor y linearly causes w.

Definition 2.1 has suggested testing for non-causality by testing
the following property:

$$(2.8) \qquad z_{n+1} \perp y_{n-p}^{n} \mid z_{n-q}^{n} + u \qquad n \geq \max(p,q)$$

for some fixed value of p and q. In general (2.1) does not imply and
is not implied by (2.8). Therefore such a test may be justified only by
maintaining some supplementary hypotheses. These may be obtained by
means of the following theorem.

Theorem 2.7. Property (2.8) is true for any p under the following
conditions:

$$(2.1) \qquad z_{n+1} \perp y_0^n \mid z_0^n + u$$

$$(2.9) \qquad z_{n+1} \perp z_0^{n-q+1} \mid y_0^n + z_{n-q}^n + u$$

$$(2.10) \qquad (y_0^n + z_{n-q}^n + u) \cap (z_0^n + u) = z_{n-q}^n + u \ .$$

Proof. (2.1) and (2.9) imply that:

7

(2.11) $\qquad (y_0^n + z_0^n + u) \, z_{n+1} \in (z_0^n + u) \cap (y_0^n + z_{n-q}^n + u)$

and, by (2.10), the l.h.s. of (2.11) belongs to $(z_{n-q}^n + u)$. i.e.:

(2.12) $\qquad (y_0^n + z_0^n + u) \, z_{n+1} \in (z_{n-q}^n + u)$

i.e.

(2.13) $\qquad z_{n+1} \perp (y_0^n + z_0^n) \mid z_{n-q}^n + u$.

Clearly (2.13) implies (2.8) for any p. $\square$

The role of theorem 2.7 may be viewed as follows. Condition (2.10) means that any linear function of $(y_i, z_j, u: 0 \le i \le n, \, n-q \le j \le n)$ that is a.s. equal to a linear function of $(z_i, u: 0 \le i \le n)$ is a.s. a function of $(z_i, u: n-q \le i \le n)$ only. This condition implies, but is not equivalent, to the following property: the only linear functions on y_0^n a.s. equal to linear functions on z_0^{n-q+1} are a.s. equal to linear functions on $(z_{n-q}^n + u)$. Condition (2.10) may be viewed as a linear form of "measureable separability" as defined in Mouchart and Rolin (1979) and may be termed "linear separability". The purpose of this condition is to avoid pathologies that could link the y-process and the z-process.

For Gaussian processes, condition (2.9) may be viewed as a Markovian condition of order q for the conditional process generating $(z_{n+1} \mid y_0^n + u)$ For general processes, condition (2.9) may be viewed as "linear" Markovian condition on the residual of the projection of z_{n+1} on $(y_0^n + u)$.

8

Theorem 2.7 suggests testing for non-causality under the maintained hypotheses (2.9) and (2.10) by testing (2.8) for some fixed (p,q). This test is generally performed by testing the significance of the coefficients of $y_n,\ldots,y_{n-p}$ in the regression of z_{n+1} on $y_n,\ldots,y_{n-p}$, $z_n,\ldots,z_{n-q},u$ (u is in this case the constant term). Note that neither (2.9) nor (2.10) involve an assumption of stationarity but stationarity implies that (2.9) may be approximately satisfied for large values of q (see, e.g., Rozanov (1967)). For autoregressive processes (of order smaller than or equal to q) condition (2.9) will be exactly satisfied, whether the process is stationary or not.

3. Sims' Non-Causality.

Sims (1972) obtained the following result.

Theorem 3.1. y does not linearly cause z if and only if

$$(3.1) \qquad \forall\, n \geq 0 \qquad z_0^\infty \perp y_n \mid z_0^n + u \,.$$

Proof. A general proof of this result was given by Hosoya (1977) but we have found the following proof both simple and insightful. From (2.2), (2.1) is equivalent to:

$$(3.2) \qquad \forall\, n \geq 0 \qquad z_0^{n+1} \perp y_0^n \mid z_0^n + u \,.$$

This implies (3.1) because, by Theorem A.10, (3.2) is equivalent to:

$$(3.3) \qquad \forall\, n \geq 0 \qquad z_0^\infty \perp y_0^n \mid z_0^n + u \,.$$

Reciprocally by Theorem A3, (3.1) implies:

$$(3.4) \qquad \forall\, n \geq 0 \,, \quad \forall\, p \geq 0 \qquad z_0^\infty \perp y_p \mid z_0^{n+p} + u$$

and this implies (3.2) by theorem A.8 and because y_0^n is generated by $(y_p \colon 0 \leq p \leq n)$. $\square$

The easiest interpretation of 3.1 may be the following: the projection of y_n onto $z_0^\infty + u$ (i.e. the best linear approximation of y_n by an element of $z_0^\infty + u$) belongs in $z_0^n + u$. In other words, only the past and the current values of z_n are relevant to explain y_n.

10

Theorem 3.1 depends crucially on linearity: in terms of independence in probability, this result would be false (See e.g. Florens and Mouchart, 1980.b).

An immediate implication of theorem A.11 is the following result.

Theorem 3.2. The following properties are equivalent:

$$(3.5) \qquad \forall\, n \geq 0 \qquad z_0^\infty \perp y_n \mid z_0^n + y_0^{n-1} + u$$

$$(3.6) \qquad \forall\, n \geq 0 \qquad z_0^\infty \perp y_n \mid z_0^n + y_0^0 + u \, .$$

Therefore, (3.5) is an alternative form of linear non-causality if the initial condition y_0 is a linear function of u. In the non-linear theory, the condition analogous to (3.5) has been introduced as a modified Sims' condition so as to obtain a condition equivalent to Granger's (see e.g., Chamberlain (1980) and Florens and Mouchart (1980 a, b and c).

Note that the properties (3.1), (3.5) and (3.6) are not modified if z_0^∞ is replaced by z_{n+1}^∞ (theorem A.2).

For practical applications or for hypothesis testing Sims' definition (3.1) can be replaced by:

$$(3.7) \qquad \forall\, n \geq q \qquad z_{n+1}^{n+p} \perp y_n \mid z_{n-q}^n + u$$

for fixed values of p and q. In general (3.1) and (3.7) are not

equivalent and further hypotheses are required to guarantee an implication between these two definitions. A theorem analogous to theorem 2.7 could be given without difficulty.

An interesting problem is the relationship between Granger's and Sims' definitions when the dimensions of the future and past of the processes (2.8 and 3.7) are fixed. It should be noted that, in general, there is no relationship between these definitions.

The following theorem shows the equivalence between (2.8) and (3.7) under a condition on the marginal process generating $(z_n)_n$.

Theorem 3.3. Under the following hypothesis:

$$(3.8) \qquad \forall\, n \geq p+q \qquad z_{n+1} \perp z_{n-(p+q)}^n \mid z_{n-q}^n + u\ .$$

The following two conditions are equivalent.

$$(3.9) \qquad \forall\, n \geq \max(p,q) \qquad z_{n+1} \perp y_{n-p}^n \mid z_q^n + u$$

$$(3.10) \qquad \forall\, n \geq \max(p,q) \qquad z_{n+1}^{n+p+1} \perp y_n \mid z_q^n + u\ .$$

Proof. a) (3.9) $\Rightarrow$ (3.10) follows from the following property, which will be proved by induction.

$$(3.11) \qquad \forall\, n \geq \max(p,q) \quad \forall\, j = 1,\ldots,p+1 \quad z_{n+1}^{n+j} \perp y_n \mid z_{n-q}^n + u$$

12

(3.9) implies (3.11) with $j = 1$. Let us assume (3.11) is true for any $j \leq p$. (3.9) gives us:

$$(3.12) \qquad z_{n+j+1} \perp y_{n+j-p}^{n+j} \mid z_{n+j-q}^{n+j} + u .$$

As $y_n \in y_{n+j-p}^{n+j}$, (3.12) implies:

$$(3.13) \qquad z_{n+j+1} \perp y_n \mid z_{n+j-q}^{n+j} + u .$$

(3.8) implies

$$(3.14) \qquad z_{n+j+1} \perp z_{n-q}^{n+j-q-1} \mid z_{n+j-q}^{n+j} + u$$

and by theorem A.3, (3.13) and (3.14) imply

$$(3.15) \qquad z_{n+j+1} \perp y_n \mid z_{n-q}^{n+q} + u$$

or equivalently

$$(3.16) \qquad z_{n+j+1} \perp y_n \mid z_{j+1}^{n+j} + z_{n-q}^{n} + u$$

and by thoerem A.3, (3.11) and (3.16) give:

$$z_{n+1}^{n+j+1} \perp y_n \mid z_{n-q}^{n} + u .$$

b) (3.10) $\Rightarrow$ (3.9) follows from the following property, also verified by induction:

(3.17) $\qquad \forall\, n \geq \max(p,q) \quad \forall\, i = 0,\ldots,p \quad z_{n+1} \perp y_{n-i}^{n} \mid z_{n-q}^{n} + u$.

(3.17) is true for $i = 0$ by (3.10). Let us assume (3.17) is satisfied for any $i \leq p-1$. From (3.10) we get:

(3.18) $\qquad z_{n-i}^{n-i+1} \perp y_{n-i-1} \mid z_{n-i-1-q}^{n-i-1} + u$.

As $z_{n-i}^{n} \subset z_{n-i}^{n-i+p}$, we get from (3.18) (by corollary A.4):

(3.19) $\qquad z_{n-i}^{n-i+p} \perp y_{n-i-1} \mid z_{n-i-1-q}^{n} + u$

(3.20) $\qquad \Rightarrow z_{n+1} \perp y_{n-i-1} \mid z_{n-i-1-q}^{n} + u$.

By theorem A.3, (3.8) and (3.20) imply:

(3.21) $\qquad z_{n+1} \perp y_{n-i-1} \mid z_{n-q}^{n} + u$.

And also by theorem A.3, (3.17) and (3.21) imply:

(3.22) $\qquad z_{n+1} \perp y_{n-i-1}^{n} \mid z_{n-q}^{n} + u$.

(3.17) is then verified and the proof is completed. $\square$

14

The marginal process generating $(z_n)_{n \geq 0}$ is (linearly) autoregressive if

(3.23) $\forall \, n \geq q \quad z_{n+1} \perp z_0^n \mid z_{n-q}^n + u$.

(3.8) implied by (3.23) but (3.8) is weaker than (3.23). Note that theorem 3.3 cannot be stated in terms of conditional independence instead of conditional orthogonality because the proof depends crucially on (see parts (i)(b) and (ii)(b) of theorem A.3) the fact that, with the same notation as in the appendix, $E_1 \perp E_2 \mid E_3$ and $E_1 \perp E_4 \mid E_3 \Rightarrow$ $E_1 \perp (E_2 + E_4) \mid E_3$. This property has no equivalent in terms of conditional independence (see Florens and Mouchart (1980.b)).

4. Haugh and Pierce's Non-Causality.

Haugh and Pierce (1977) have suggested analysing the cross-correlations between the innovations of the z-process and the innovations of the y-process. In this section we compare the approach of Haugh and Pierce and linear non-causality.

In our notation, the innovations of a process $(y_n)_{n \geq 0}$ form the process denoted $y_n - y_0^{n-1} y_n = (y_0^{n-1})^\perp y_n$ i.e. the difference between y_n and its projection on the linear space of all the linear combinations of $y_0, \ldots, y_{n-1}$, or, alternatively, the projection of y_n on the orthogonal complement of this space of linear combinations. So the property stated in the following theorem can be viewed as the Haugh and Pierce definition of non-causality, rewritten in our notation.

Theorem 4.1. If y does not linearly cause z, then

$$(4.1) \qquad \begin{array}{l} \forall\, n \geq 0 \\ \forall\, p < n \end{array} \qquad (z_{n+1} \mid z_0^n + u) \perp (y_{p+1} \mid y_0^p) \, .$$

Proof. Clearly $(y_0^p)^\perp y_{p+1} \subset y_0^n$ if $p < n$. Therefore, condition (2.1) implies:

$$(4.2) \qquad \forall\, n \geq 0 \, , \, \forall\, p < n : z_{n+1} \perp (y_0^p)^\perp y_{p+1} \mid z_0^n + u \, . \quad \square$$

Theorem 4.2. If $y_0 \in z_0^n + u$, then (4.1) implies that y does not linearly cause z.

Proof. We first rewrite (4.1) as follows:

$$(4.3) \qquad \begin{array}{l} \forall \, n \geq 0 \\ \forall \, p < n \end{array} \qquad (z_0^n + u)^{\perp} z_{n+1} \perp y_{p+1} \mid y_0^p \, .$$

By theorem A.9, (4.3) is equivalent to:

$$(4.4) \qquad \begin{array}{l} \forall \, n \geq 0 \\ \forall \, p < n \end{array} \qquad (z_0^n + u)^{\perp} z_{n+1} \perp y_{p+1} \mid y_0^0$$

i.e.

$$(4.5) \qquad \begin{array}{l} \forall \, n \geq 0 \\ \forall \, p < n \end{array} \qquad (z_{n+1} \mid z_0^n + u) \perp (y_{p+1} \mid y_0^0) \, ,$$

and, by theorem A.8, (4.5) is equivalent to

$$(4.6) \qquad \forall \, n \geq 0 \qquad (z_{n+1} \mid z_0^n + u) \perp (y_0^n \mid y_0^0)$$

and (4.6) is equivalent to (2.1), by theorem A.6, if $y_0 \in z_0^n + u$. $\square$

These theorems show that the equivalence between Haugh and Pierce's condition (4.1) and linear non-causality basically depends on the specification of the initial condition y_0. If u has the form $u = y_0 + v$, then the two approaches are equivalent; otherwise linear non-causality implies, but is not implied, by condition (4.1).

5. Rational Expectations and Non-Causality.

Non-causality may be viewed as the condition that the prediction of z_{n+1} based only on its own history z_0^n will not improve if it is also based on the past history of y. This suggests that "y does not cause z" may be rephrased as "z is self-predictive w.r.t. y"(for more justification see e.g. Florens and Mouchart (1980c)).

Consider now a sequence of messages and, associated with it, a sequence of "information sets" I_n representing the information contained in all the messages up to instant n. Then one may decompose z_n (or y_n) into an "expected" component given I_n and an "unexpected" one. An interesting question is to analyze non-causality in terms of such a decomposition for both y and z. This was the object of a recent paper by Sims (1980).

In a linear context, the sequence I_n will be an increasing sequence of closed subspaces of L^2; often I_n will have the form w_0^n where $\{w_n\}$ is a sequence of "observations". The "expected" component $\hat{z}_n$ of z_n becomes the projection of z_n on I_n, and its "unexpected" component ϵ_n becomes the projection of z_n on the orthogonal complement of I_n. In other words, we have the following decomposition:

$$(5.1) \qquad z_n = \hat{z}_n + \epsilon_n \qquad \begin{aligned} \hat{z}_n &= I_n z_n \\ \epsilon_n &= I_n^{\perp} z_n \end{aligned} \quad .$$

Similarly for y_n:

$$(5.2) \qquad y_n = \hat{y}_n + \eta_n \qquad \begin{aligned} \hat{y}_n &= I_n \, y_n \\ \eta_n &= I_n^{\perp} \, y_n \end{aligned} \quad .$$

Remark. It should be pointed out that a closed subspace is a poor mathematical translation of the intuitive concept of information. Indeed, "knowing" w_n should involve knowing all (measurable) transformations of w_n, and not only the linear ones. In other words, the σ-field is the natural translation for the concept of information. In the present context, representing expectations by projection on a space of linear functions (defined on w_0^n) only may be justified on the grounds of computational simplicity or by a Gaussian assumption. Finally, the natural way to deal with an increasing sequence of σ-fields is to intro-duce a filtration and to consider as the rational expectation of a given process the nearest process adapted to that filtration (see, e.g. Dellacherie-Meyer, (1976), see also Futia (1981)).

In Sims (1980), the information sets I_n are taken to be:

$$(5.3) \qquad I_n = z_0^{n-1} + y_0^{n-1} + u \qquad n \geq 1 \; .$$

Therefore:

$$(5.4) \qquad \varepsilon_n = (z_0^{n-1} + y_0^{n-1} + u)^{\perp} z_n \qquad n \geq 1$$

$$(5.5) \qquad \eta_n = (z_0^{n-1} + y_0^{n-1} + u)^{\perp} y_n \qquad n \geq 1 \; .$$

It is also assumed that the initial conditions z_0 and y_0 are totally unexpected i.e.

(5.6) $\qquad$ $\epsilon_0 = z_0$ and $\eta_0 = y_0$.

With the above notation, we have the following results.

$\quad$ Theorem 5.1. If y does not linearly cause z, then:

(5.7) $\qquad$ $\forall\, n \geq 0 \qquad z_{n+1} \perp \eta_0^n \mid \epsilon_0^n + u$.

$\quad$ Proof. Step 1. We first prove:

(5.8) $\qquad$ $\forall\, n \geq 1 \qquad \epsilon_n = (z_0^{n-1} + u)^{\perp} z_n$.

We note that condition (2.1) may be written as:

(5.9) $\qquad$ $\forall\, n \geq 1 \quad (y_0^{n-1} + z_0^{n-1} + u)\, z_n = (z_0^{n-1} + u)\, z_n$.

This is equivalent to (5.8).

$\quad$ Step 2. We now prove:

(5.10) $\qquad$ $\epsilon_0^n + u = z_0^n + u \qquad n \geq 0$.

Given (5.6), this is trivial for $n = 0$. Suppose (5.10) is true for some n. Then from (5.8), we have

$$\epsilon_{n+1} = z_{n+1} - (z_0^n + u)\, z_{n+1}$$

which clearly belongs to $z_0^{n+1} + u$.

20

Therefore, under (5.10) for some n, $\varepsilon_0^{n+1} + u \subset z_0^{n+1} + u$. Conversely, from (5.9) again, $z_{n+1} = \varepsilon_{n+1} + (z_0^n + u) z_{n+1}$ which, under (5.10) for some n, belongs to $\varepsilon_0^{n+1} + u$; therefore under (5.10) for some n, $\varepsilon_0^{n+1} + u \supset z_0^{n+1} + u$.

Step 3. We now prove:

$$(5.11) \qquad \eta_0^n \subset y_0^n + z_0^n + u \qquad \forall\, n \geq 0$$

indeed,

$$\eta_p = y_p - (y_0^{p-1} + z_0^{p-1} + u)\, y_p \in y_0^n + z_0^n + u \qquad \forall\, p \leq n .$$

The proof is concluded by noticing that (2.1) may be written as:

$$(5.12) \qquad z_{n+1} \perp (y_0^n + z_0^n + u) \mid z_0^n + u . \qquad \Box$$

Theorem 5.2. If $y_0 \in u$, then condition (5.7) implies that y does not linearly cause z.

Proof. Step 1. We first prove that the assumptions imply:

$$(5.13) \qquad \forall\, n \geq 0 \qquad z_{n+1} \perp y_0^n \mid \varepsilon_0^n + u ,$$

indeed, (5.7) is equivalent to:

$$(\epsilon_0^n + u)^{\perp} z_{n+1} \perp \eta_0^n \qquad\qquad \forall\, n \geq 0$$

$$\Leftrightarrow (\epsilon_0^n + u)^{\perp} z_{n+1} \perp \eta_p \qquad\qquad \forall\, (p,n)\ \ 0 \leq p \leq n$$

$$\Leftrightarrow (\epsilon_0^n + u)^{\perp} z_{n+1} \perp (y_0^{p-1} + z_0^{p-1} + u)^{\perp} y_p \quad \forall\, (p,n)\ \ 0 \leq p \leq n$$

$$\Leftrightarrow (\epsilon_0^n + u)^{\perp} z_{n+1} \perp y_p \mid y_0^{p-1} + z_0^{p-1} + u \quad \forall\, (p,n)\ \ 0 \leq p \leq n$$

$$\Rightarrow (\epsilon_0^n + u)^{\perp} z_{n+1} \perp y_p \mid y_0^0 + z_0^0 + u \quad (\text{theorem A.9})$$

$$\forall\, (p,n)\ \ 0 \leq p \leq n$$

$$\Leftrightarrow (\epsilon_0^n + u)^{\perp} z_{n+1} \perp y_p \mid z_0^0 + u \quad (\text{because } y_0 \in u)$$

$$\forall\, (p,n)\ \ 0 \leq p \leq n$$

$$\Rightarrow z_{n+1} \perp y_p \mid \epsilon_0^n + u \qquad\qquad \forall\, (p,n)\ \ 0 \leq p \leq n\ .$$

The last step is made by using the fact that $z_0^0 + u \subset \epsilon_0^n + u$ and by applying theorem A.6.

Step 2. We now prove (2.1) by induction. It is clearly true for n = 0 under the hypothesis $y_0 \in u$. Now suppose (2.1) is true for $n \leq p$; we prove that (2.1) is also true for n = p+1. From step 1 in theorem 5.1, (5.8) is true for $n \leq p$ and, from step 2, (5.10) is also true for $n \leq p$; in particular $\epsilon_0^p + u = z_0^p + u$; therefore, by (5.14), (2.1) is true for n = p+1. $\square$

Theorem 5.1 shows that condition (5.7) is, in general, stronger than non-causality but theorems 5.2 shows that condition (5.7) is actually equivalent to non-causality if the initial condition on y is included in u (which implies that both expectations $\hat{z}_n$ and $\hat{y}_n$ may involve y_0.)

Extensions of non-causality properties can be easily done in terms of expected or unexpected components of the variables. As an example we give the following version of theorem 3.1.

Theorem 5.3. If $y_0 \in u$, y does not linearly cause z if and only if

$$(5.14) \qquad \forall\, n \geq 0 \qquad z_{n+1}^{\infty} \perp \eta_0^n \mid \varepsilon_0^n + u \ .$$

Proof. Using theorems 5.1 and 5.2 we just have to prove the equivalence of (5.14) and (5.7).

(5.14) implies (5.7) by theorem A.2 (since $z_{n+1} \in z_{n+1}^{\infty}$). (5.7) implies (5.14) by using theorem A.10 (note that if y does not linearly cause z, we have $\forall\, n \quad z_0^n + u = \varepsilon_0^n + u$ – see step 2 of the proof of theorem 5.1). □

Finally let us note that the property "y does not cause z" implies the following conditional orthogonality:

$$(5.15) \qquad \forall\, n \geq 0 \quad \forall\, k \geq 0 \qquad z_{n+k+1} \perp \eta_{n+1}^{n+k} \mid \varepsilon_0^{n+k} + \eta_0^n + u$$

(by theorem 5.1 and corollary A.4). This property is actually the property tested by Sims in his 1980 paper.

Appendix. Orthogonality.

Let $X = (L, <\cdot,\cdot>)$ be a Hilbert space on $\mathbb{R}$, i.e. L is a linear space of vectors x_i's and $<\cdot,\cdot>$ is an inner product (bilinear, symmetric and positive definite) which makes L complete. Details on this structure can be found in e.g. Halmos (1957) or Greub (1975). The Hilbert space we use in the main body of this paper is the set of (classes of) random variables y defined (up to an almost sure equality) on a probability space (Ω, G, P) and such that $E(y^2)$ is finite. The inner product between y and z is then defined by $E(yz)$. However, definitions and results given in this appendix are stated in terms of a general Hilbert space.

Let E_i be complete (or, equivalently, closed) linear subspaces of L; where in particular E_0 is the subspace containing the null vector only. $E_i + E_j$ denotes the usual sum of subspaces, $x_1 \perp x_2$ denotes the usual orthogonality w.r.t. $<\cdot,\cdot>$ (i.e. $<x_1, x_2> = 0$). Likewise $x \perp E$ means $x \perp e$ $\forall e \in E$ and $E_1 \perp E_2$ means $e_1 \perp e_2$ $\forall e_1 \in E_1$, $e_2 \in E_2$. $E^\perp$ denotes the orthogonal complement of E, i.e. $E^\perp = \{x \in L | x \perp E\}$. Note that $E_0^\perp = L$. Finally, Ex denotes the orthogonal projection of the vector x on the subspace E, i.e. the unique vector $e \in E$ such that $(x-e) \perp E$. Note that this is a linear idempotent operation. We shall make use of the following property:

$$<x_1, Ex_2> = <Ex_1, x_2> = <Ex_1, Ex_2> \quad \forall x_1, x_2 \in L \text{ and } \forall E .$$

Note that $Lx = x$ and $E_0 x = 0$. This notation is extended as follows: $E_2 E_1$ means the projection of E_1 on E_2. Note that $E_0 E_1 = E_0$ and $L E_1 = E_1$. We shall also use:

24

$$E_1 \subset E_2 \Rightarrow E_1 E_2 = E_2 E_1 = E_1.$$

We shall introduce in this appendix the two concepts of conditional orthogonality and biconditional orthogonality. They are not new concepts, as they are merely particular cases of orthogonality, but they provide convenient notation and results for our kind of problems.

A.1. <u>Theorem</u>.

The following properties are equivalent and define "E_1 and E_2 are orthogonal conditionally on E_3" which is denoted as "$E_1 \perp E_2 \mid E_3$":

(i) $\qquad (x_1, x_2) \in E_1 \times E_2 \Rightarrow (x_1 - E_3 x_1) \perp (x_2 - E_3 x_2)$

$\qquad$ (or $E_3^{\perp} E_1 \perp E_3^{\perp} E_2$ or $E_3^{\perp} E_1 \perp E_2$ or $E_1 \perp E_3^{\perp} E_2$)

(ii) $\qquad x_1 \in E_1 \Rightarrow (E_2 + E_3) x_1 = E_3 x_1$

$\qquad$ (or $(E_2 + E_3) E_1 = E_3 E_1$)

(iii) $\qquad x_3 \in E_2 \Rightarrow (E_1 + E_3) x_2 = E_3 x_2$

$\qquad$ (or $(E_1 + E_3) E_2 = E_3 E_2$) . $\quad \square$

From (i), conditional orthogonality may be interpreted as the usual orthogonality between the projections of E_1 and E_2 on $E_3^{\perp}$. In a statistical context, these projections will be recognized as the residuals in regression analysis.

A.2. Elementary properties.

(i) $\quad\quad E_1 \perp E_2 \mid E_0 \Leftrightarrow E_1 \perp E_2$

(ii) $\quad\quad E_4 \subset E_1$ and $E_1 \perp E_2 \mid E_3$ imply $E_4 \perp E_2 \mid E_3$

(iii) $\quad\quad E_4 \subset E_3$ implies $E_1 \perp E_4 \mid E_3 \quad\quad \forall E_1$

(iv) $\quad\quad E_4 \subset E_3$ and $E_1 \perp E_2 \mid E_3$ imply $(E_1 + E_4) \perp E_2 \mid E_3$.

A.3. Fundamental property of conditional orthogonality.

The following properties are equivalent:

(i) $\quad\quad$ (a) $E_1 \perp E_2 \mid E_3$ and (b) $E_1 \perp E_4 \mid E_3$

(ii) $\quad\quad$ (a) $E_1 \perp E_2 \mid E_3$ and (b) $E_1 \perp E_4 \mid E_2 + E_3$

(iii) $\quad\quad E_1 \perp (E_2 + E_4) \mid E_3$.

Proof. (iii) $\Rightarrow$ (i) by A.2 iii. To prove that (i) $\Rightarrow$ (iii) take $x \in E_2 + E_4$, i.e. $x = x_2 + x_4$. Then, by the linearlity of projection (i) implies that $(E_1 + E_3) x = E_3 x$. Similarly, (iii) $\Rightarrow$ (ii a) by A.2 iii and (iii) $\Rightarrow$ (ii b) because, by (iii), $(E_2 + E_3 + E_4) x_1 = E_3 x_1 \quad \forall x_1 \in E_1$ which implies $(E_2 + E_3) x_1 = E_3 x_1$ as a property of projections. Finally (ii) $\Rightarrow$ (iii) because for any $x_1 \in E_1$, $(E_2 + E_3 + E_4) x_1 = (E_2 + E_3) x_1 = E_3 x_1$ by (ii b) and (ii a) successively. $\square$

A.4. **Corollary**.

$$E_4 \subset E_1 + E_3 \quad \text{and} \quad E_1 \perp E_2 \mid E_3 \quad \text{imply} \quad E_1 \perp E_2 \mid E_4 + E_3 \, . \quad \square$$

A.5. **Theorem**.

The following properties are equivalent and define "E_1 conditionally on E_2 and E_3 conditionally on E_4 are orthogonal" which is denoted by "$(E_1 \mid E_2) \perp (E_3 \mid E_4)$":

(i) $\forall \, (x_1, x_3) \in E_1 \times E_3 \Rightarrow (x_1 - E_2 x_1) \perp (x_3 - E_4 x_3)$

(ii) $E_2^{\perp} E_1 \perp E_4^{\perp} E_3$

(iii) $E_1 \perp E_4^{\perp} E_3 \mid E_2$

(iv) $E_2^{\perp} E_1 \perp E_3 \mid E_4 \, .$

A.6. **Elementary properties**.

(i) $E_1 \perp E_2 \mid E_3 \Leftrightarrow (E_1 \mid E_3) \perp (E_2 \mid E_3)$

       in particular: $E_1 \perp E_2 \Leftrightarrow (E_1 \mid E_0) \perp (E_2 \mid E_0)$

(ii) $E_1 \perp E_2 \mid E_3 \Leftrightarrow (E_1 \mid E_0) \perp (E_2 \mid E_3)$

            $\Leftrightarrow E_1 \perp (E_2 \mid E_3)$

(iii) $E_5 \subset E_1 \quad \text{and} \quad (E_1 \mid E_2) \perp (E_3 \mid E_4) \quad \text{imply} \quad (E_5 \mid E_2) \perp (E_3 \mid E_4)$

(iv) $E_4 \subset E_2 \quad \text{and} \quad (E_1 \mid E_2) \perp (E_3 \mid E_4) \Leftrightarrow E_1 \perp E_3 \mid E_2 \, .$

A.7. <u>Fundamental property of biconditional orthogonality.</u>

The following properties are equivalent:

(i)   (a) $(E_1 \mid E_2) \perp (E_3 \mid E_4)$ and (b) $(E_1 \mid E_2) \perp (E_5 \mid E_4)$

(ii) (a) $(E_1 \mid E_2) \perp (E_3 \mid E_4)$ and (b) $(E_1 \mid E_2) \perp (E_5 \mid E_3 + E_4)$

(iii) $(E_1 \mid E_2) \perp [(E_3 + E_5) \mid E_4]$.

   <u>Proof</u>. From theorem A.5 (iv) one may replace $(E_1 \mid E_2)$ by
$E_2^{\perp} E_1$ and then use A.3. □

   A.8. <u>Theorem</u>.

   Let A be a subset of L and E be the closed linear subspace
of L generated by A. Then for any closed linear subspaces E_1, E_2, E_3
the following two properties are equivalent:

(i) $A \mid E_1 \perp E_2 \mid E_3$

(ii) $E \mid E_1 \perp E_2 \mid E_3$.

   <u>Sequences of conditional orthogonalities</u>.

   A.9. <u>Theorem</u>.

   Let $(F_n)_{n \geq 0}$ be an increasing sequence $(F_{n-1} \subset F_n)$ of closed
linear subspaces of L. Then, for any E and G, the following pro-
perties are equivalent:

(i) $\forall\, n \geq 0$ $E \perp F_n \mid F_{n-1} + G$

(ii) $\forall\, n \geq 0$ $E \perp F_n \mid F_0 + G$.

Proof. (ii) implies (i) by A.4. The converse follows from the property: $\forall\, n \geq 0,\ \forall\, q = 0,\ldots,n$ $E \perp F_n \mid F_{n-q}+G$ which is proved by induction. It is clearly true for $q = 1$. Let us assume this property for general q and note that (i) implies $E \perp F_{n-q} \mid F_{n-(q+1)}+G$. The result follows from application of A.3 and A.4. $\square$

A.10. **Theorem.**

Let $(E_n)_{n \geq 0}$ and $(F_n)_{n \geq 0}$ be increasing sequences of closed linear subspaces of L such that $\forall\, n$ $F_n \subset E_n$. E_∞ is the closed subspace generated by $\bigcup_{n \geq 0} E_n$. Then the following properties are equivalent:

(i) $\forall\, n \geq 0$ $E_{n+1} \perp F_n \mid E_n + G$

(ii) $\forall\, n \geq 0$ $E_\infty \perp F_n \mid E_n + G$.

Proof. (ii) implies (i) by A.2. The converse follows by using A.6 and the property:

$$\forall\, n \geq 0 \quad \forall\, p \geq 1 \qquad E_{n+p} \perp E_n \mid E_n + G .$$

This property is proved by induction. It is true for $p = 1$. Let us assume this property is true for general p. (i) implies $E_{n+p+1} \perp F_{n+p} \mid E_{n+p} + G$ and A.3 and A.4 implies the result. $\square$

A.11. **Theorem**.

With the same definitions as in the preceding theorem the following two properties are equivalent:

(i) $E_\infty \perp F_n \mid E_n + F_{n-1} + G$

(ii) $\forall\, n \geq 0$ $E_\infty \perp F_n \mid E_n + F_0 + G.$

The proof is essentially the same as the proof of theorem A.6.

References

Anderson, T.W. (1971), The Statistical Analysis of Time Series, Wiley, New York.

Bahadur, R.R. (1954), Sufficiency and Statistican Decision Functions, The Annals of Mathematical Statistics, 25, 423-462.

Chamberlain, G. (1980), On the General Non-Equivalence of Granger and Sims Exogeneity, unpublished University of Wisconsin manuscript.

Dellacherie, C., and Meyer, P.A. (1976), Probabilite et Potentiel, Hermann, Paris.

Engle, R., Hendry, D. and Richard, J.F. (1980), Exogenity, Causality and Structural Invariance in Econometric Modeling, CORE D.P. 8083, University Catholique de louvain.

Florens, J.P., and Mouchart, M. (1980a), Initial and Sequential Reduction of Bayesian Experiments, CORE D.P. 8015, Universite Catholique de louvain.

Florens, J.P., and Mouchart, M. (1980b), A Note on Non-Causality, to appear in Econometrica.

Florens, J.P., and Mouchart, M. (1980c), Conditioning in Econometric Models, CORE D.P. 8042, Universite Catholique de louvain

Florens, J.P., Mouchart, M., and Rolin, J.M. (1980), Reduction dans les experiences bayesiennes sequentielles, Cahiers du Centre d'etudes de recherche operationnelle, 22, No. 3-4, 353-362.

Futia, C. (1981), Rational Expectations in Stationary Linear Models, Econometrica, 49.1, 171-192.

Gourieroux, C., and Montfort, A. (1980), Sufficient Linear Structures: Econometric Applications, Econometrica, 48.5, 1083-1098.

Granger, C.W.J. (1969), Investigating Causal Relations by Econometric
Models and Cross Spectral Methods, Econometrica, 37, 424-438.

Greub, (1975), Linear Algebra, 4th ed., Springer Verlag, New York.

Hall, W.J., Wijsman, R.A., and Ghosh, J.K. (1965), The Relationship
Between Sufficiency and Invariance with Applications in Sequential
Analysis, The Annals of Mathematical Statistics, 36, 575-614.

Halmos, P. (1957), Introduction to Hilbert Spaces and the Theory of
Spectral Multiplicity, 2nd ed., Chelsea Pub., New York.

Haugh, L.D. and Pierce, D.A. (1977), Causality in Temporal Systems:
Characterizations and a Survey, Journal of Econometrics, 5,
265-293.

Hosoya, Y. (1977), On the Granger Condition for Non-Causality, Econometrica,
45, 1735-1736.

Mouchart, M., and Rolin, J.M. (1979), On k-Sufficiency, CORE D.P. 7925,
University Catholique de Louvain.

Neveu, J. (1964), Bases mathematiques du calcul des probabilites, Masson,
Paris.

Rozanov, Y.A. (1967), Stationary Random Processes, Holden Day, San
Francisco.

Sims, C.A. (1972), Money, Income and Causality, American Economic Review,
62, 540-562.

Sims, C.A. (1980), Manoeconomics and Reality, Econometrica, 49, 171-192.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.

2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.

3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.

4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.

5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.

6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.

7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.

8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.

9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.

10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.

11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.

12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.

13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.

14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.

15. "Maximum Likelihood Estimation of Parameter of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.

16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time series," Steffen L. Lauritzen, April, 1974.

17. "General Exponential Models for Discrete Observations,"
    Steffen L. Lauritzen, May, 1974.

18. "On the Interrelationships among Sufficiency, Total Sufficiency and
    Some Related Concepts," Steffen L. Lauritzen, June, 1974.

19. "Statistical Inference for Multiply Truncated Power Series Distributions,"
    T. Cacoullos, September 30, 1974.

    Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models
    in the Time and Frequency Domains," T. W. Anderson, June 1975.

21. "Asymptotic Properties of Some Estimators in Moving Average Models,"
    Raul Pedro Mentz, September 8, 1975.

22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting,"
    Mituaki Huzii, February 1976.

23. "Estimating Means when Some Observations are Classified by Linear
    Discriminant Function," Chien-Pai Han, April 1976.

24. "Panels and Time Series Analysis: Markov Chains and Autoregressive
    Processes," T. W. Anderson, July 1976.

25. "Repeated Measurements on Autoregressive Processes," T. W. Anderson,
    September 1976.

26. "The Recurrence Classification of Risk and Storage Processes,"
    J. Michael Harrison and Sidney I. Resnick, September 1976.

27. "The Generalized Variance of a Stationary Autoregressive Process,"
    T. W. Anderson and Raul P. Mentz, October 1976.

28. "Estimation of the Parameters of Finite Location and Scale Mixtures,"
    Javad Behboodian, October 1976.

29. "Identification of Parameters by the Distribution of a Maximum
    Random Variable," T. W. Anderson and S.G. Ghurye, November 1976.

30. "Discrimination Between Stationary Guassian Processes, Large Sample
    Results," Will Gersch, January 1977.

31. "Principal Components in the Nonnormal Case: The Test for Sphericity,"
    Christine M. Waternaux, October 1977.

32. "Nonnegative Definiteness of the Estimated Dispersion Matrix in a
    Multivariate Linear Model," F. Pukelsheim and George P.H. Styan, May 1978.

TECHNICAL REPORTS (continued)

33. "Canonical Correlations with Respect to a Complex Structure,"
    Steen A. Andersson, July 1978.

34. "An Extremal Problem for Positive Definite Matrices," T.W. Anderson and
    I. Olkin, July 1978.

35. " Maximum likelihood Estimation for Vector Autoregressive Moving
    Average Models," T. W. Anderson, July 1978.

36. "Maximum likelihood Estimation of the Covariances of the Vector Moving
    Average Models in the Time and Frequency Domains," F. Ahrabi, August 1978.

37. "Efficient Estimation of a Model with an Autoregressive Signal with
    White Noise," Y. Hosoya, March 1979.

38. "Maximum Likelihood Estimation of the Parameters of a Multivariate
    Normal Distribution,"T.W. Anderson and I. Olkin, July 1979.

39. "Maximum Likelihood Estimation of the Autoregressive Coefficients and
    Moving Average Covariances of Vector Autoregressive Moving Average Models,"
    Fereydoon Ahrabi, August 1979.

40. "Smoothness Priors and the Distributed Lag Estimator," Hirotugu Akaike,
    August, 1979.

41. "Approximating Conditional Moments of the Multivariate Normal Distribution,"
    Joseph G. Deken, December 1979.

42. "Methods and Applications of Time Series Analysis – Part I: Regression,
    Trends, Smoothing, and Differencing," T.W. Anderson and N.D. Singpurwalla,
    July 1980.

43. "Cochran's Theorem, Rank Additivity, and Tripotent Matrices." T.W. Anderson
    and George P.H. Styan, August, 1980.

44. "On Generalizations of Cochran's Theorem and Projection Matrices,"
    Akimichi Takemura, August, 1980.

45. "Existence of Maximum Likelihood Estimators in Autoregressive and
    Moving Average Models," T.W. Anderson and Raúl P. Mentz, Oct. 1980.

46. "Generalized Correlations in the Singular Case," Ashis Sen Gupta,
    November 1980.

47. "Updating a Discriminant Function on the Basis of Unclassified Data,"
    G.J. McLachlan and S. Ganesalingam, November 1980.

48. "A New Proof of Admissibility of Tests in the Multivariate Analysis
    of Variance." T.W. Anderson and Akimichi Takemura, January, 1981.

49 "A-Optimality for Regression Designs," N. N. Chan, January 1981.

50. "Tests for Standardized Generalized Variances of Multivariate Normal
    Populations of Possibly Different Dimensions," Ashis Sen Gupta,
    February 1981.

51. "A Linear Theory for Noncausality," J.P. Florens and M. Mouchart,
    November 1981.

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
Technical Report No. 51	AD-A115302	

4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED
A LINEAR THEORY FOR NONCAUSALITY	Technical Report
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)
J. P. Florens and M. Mouchart	N00014-75-C-0442

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Department of Statistics, Sequoia Hall Stanford University Stanford, CA	(NR-042-034)

11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE
Office of Naval Research Statistics and Probability Program Code 436 Arlington, Virginia 22217	November 1981
	13. NUMBER OF PAGES
	32

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report)
	UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Stochastic processes; noncausality; projections in Hilbert spaces; innovations and rational expectations.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Different definitions of noncausality (according to Granger, Sims, Haugh and Pierce,...) are analyzed in terms of orthogonality in the Hilbert space of square integrable variables. Conditions, when necessary, are given for their respective equivalence. Some problems of testability are mentioned. Finally noncausality is also analyzed in terms of "rational expectations," extending previous results of Sims.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73
S/N 0102-014-6601

DATE
FILME
7—8